



# NMR metabolomic approaches for plants, toxicology and medicine

Cécile Canlet, Laetitia L. Shintu, Benedicte B. Elena-Herrmann

## ► To cite this version:

Cécile Canlet, Laetitia L. Shintu, Benedicte B. Elena-Herrmann. NMR metabolomic approaches for plants, toxicology and medicine. *L'Actualité Chimique*, 2012, 364, pp.59-68. hal-00939569

**HAL Id: hal-00939569**

**<https://hal.science/hal-00939569>**

Submitted on 30 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# La métabolomique par RMN et ses applications en biologie végétale, toxicologie et médecine

Cécile Canlet, Laetitia Shintu et Bénédicte Elena-Herrmann

- Résumé** Cet article présente les principes ainsi que les outils analytiques et chimiométriques qui sont au cœur du développement des approches d'analyse métabolomique par RMN. Il en illustre les applications variées au travers d'exemples choisis dans différents domaines de la biologie et de la médecine, concernant l'étude des organismes génétiquement modifiés, les interactions plantes/environnement, la toxicologie environnementale, la pharmacologie, la cancérologie ou l'épidémiologie moléculaire.
- Mots-clés** **RMN, métabolomique, métabolites, chimiométrie, biostatistique, biofluides.**
- Abstract** **NMR metabolomic approaches for plants, toxicology and medicine**  
This article describes the principles as well as the analytical and chemometric tools that support the development of NMR metabolomic approaches. These analytical methods are illustrated through a range of selected examples in various domains of biology and medicine, concerning the study of genetically modified organisms, plant/environment interactions, environmental toxicology, pharmacology, cancerology, or molecular epidemiology.
- Keywords** **NMR, metabolomics, metabolites, chemometrics, biostatistics, biofluids.**

Si l'essor de la technique de RMN a été lié pendant plusieurs décennies au développement de méthodes pour l'étude de systèmes moléculaires de plus en plus complexes, tels que protéines ou assemblages macromoléculaires, l'étude par RMN de petites molécules dans des systèmes naturels fait l'objet d'un intérêt ravivé et croissant depuis une quinzaine d'années. À l'opposé des problématiques classiques de détermination structurale sur des produits isolés et souvent isotopiquement marqués, l'enjeu réside à présent dans la caractérisation de mélanges complexes, impliquant la détection et l'analyse simultanée d'un grand nombre de molécules individuellement simples. Telle est l'ambition d'une analyse de type *métabolomique*, dont l'objet se porte généralement sur l'étude de fluides biologiques (urine, sérum ou plasma sanguin, liquide amniotique, liquide céphalo-rachidien...), d'extraits de plantes, de cellules, par RMN en solution, mais également sur l'analyse de tissus intacts (biopsies) ou d'organismes modèles entiers (levures, nématodes *C. elegans*...) par la technique de RMN dite « à haute résolution à l'angle magique » (HR-MAS) adaptée pour des échantillons semi-solides. La métabolomique peut être définie comme l'identification et la quantification non sélective de l'ensemble des *métabolites* d'un système biologique (voir encadré 1).

Le développement de ces approches a été nettement favorisé par les progrès technologiques récents de la RMN, notamment l'accessibilité des hauts champs magnétiques (600 MHz et plus) qui améliore simultanément sensibilité et résolution spectrale, mais également la mise au point de technologies robustes pour l'acquisition de données à haut débit (jusqu'à une centaine d'échantillons par jour) permettant l'analyse de grandes cohortes d'échantillons [1]. Les travaux de J. Lindon et J.K. Nicholson (Imperial College,

Londres) ont été pionniers dans l'exploitation de ces ressources analytiques pour la détection d'empreintes métaboliques, contribuant de façon majeure aux développements des protocoles expérimentaux et des méthodes adaptées pour l'exploitation des données spectrales, tout en dévoilant le vaste potentiel des approches de métabolomique pour des applications dans de nombreux domaines, tels que le diagnostic médical, la toxicologie, l'épidémiologie moléculaire, ou plus récemment de la médecine personnalisée.

La RMN contribue sans exclusivité, parmi un ensemble de techniques analytiques, à l'essor de cette vaste entreprise d'exploration du métabolome, qui inclut notamment des approches de spectrométrie infrarouge, des techniques de

## Encadré 1

### Des métabolites au métabolisme : quelques définitions

**Métabolisme** : ensemble de toutes les réactions chimiques présentes dans un organisme et dont les réactifs, intermédiaires et produits sont les métabolites.

**Métabolite** : petite molécule présente dans un organisme, de masse molaire inférieure à 1 kDa (1 kg/mol). Les métabolites dits *primaires* sont directement impliqués dans les processus normaux de croissance, développement ou reproduction d'un organisme, et comprennent entre autres acides aminés, lipides, sucres ou acides nucléiques. Les métabolites *secondaires* ne sont pas essentiels à ces processus et sont associés généralement au métabolisme des plantes, bactéries et champignons. Ils participent principalement à l'adaptation de ces organismes à leur environnement.

**Métabolome** : ensemble des métabolites présents dans un organisme.



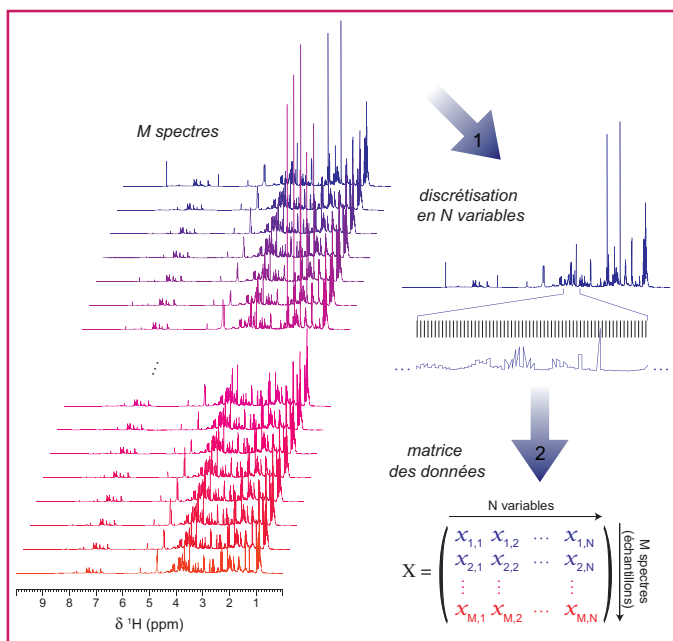


Figure 2 - Principe de mise en forme d'un jeu de données RMN 1D pour une cohorte de M échantillons.

Les spectres sont segmentés en N régions de largeur fixe, par exemple  $10^{-3}$  ppm, ou éventuellement variable. L'intégrale  $x_{i,j}$  du signal RMN mesurée pour chaque segment produit une matrice X, de dimensions (M, N), synthétisant l'ensemble de l'information spectrale exploitable pour la cohorte d'échantillons étudiée.

d'échantillons analysés. D'autres procédés de transformation des données existent, tels que des méthodes de déconvolution qui permettent une estimation directe de la

concentration relative ou absolue des métabolites par reconnaissance de leur enveloppe spectrale et intégration totale des signaux correspondants. Elles restent toutefois peu exploitées car il n'existe pas à ce jour d'outils informatiques pour leur application fiable et automatisée à l'analyse de larges cohortes de spectres RMN.

L'analyse métabolomique privilégie ensuite la modélisation des données RMN ainsi formatées par des approches d'analyse statistique multivariée (voir encadré 3). L'intérêt de ces méthodes est de réduire la dimension de l'espace des variables (initialement N) en recherchant les combinaisons linéaires de celles-ci, appelées composantes ou variables latentes, qui synthétisent l'information métabolique recherchée. Les résultats des analyses multivariées peuvent être visualisés dans deux espaces distincts (figure 3) : l'espace des individus ou « scores », dans lequel les échantillons s'arrangent en sous-populations (ou classes), et l'espace des variables ou « loadings », qui met en évidence les *signatures métaboliques* correspondantes, c'est-à-dire les différences caractéristiques, entre classes d'échantillons, de concentration des métabolites détectés.

Deux grandes familles de méthodes statistiques multivariées se distinguent selon le caractère *non supervisé* ou *supervisé* de l'analyse. La première s'intéresse à la description des caractéristiques intrinsèques des profils métaboliques définissant un jeu de données, sans connaissance *a priori* de groupes d'appartenance des échantillons, tandis que les méthodes supervisées exploitent des informations complémentaires sur la nature des échantillons (par exemple individus sains vs. malades, traités vs. non traités, etc.). La méthode non supervisée la plus répandue est l'analyse en composantes principales (ACP),

### Encadré 3

#### Principes d'analyse statistique multivariée

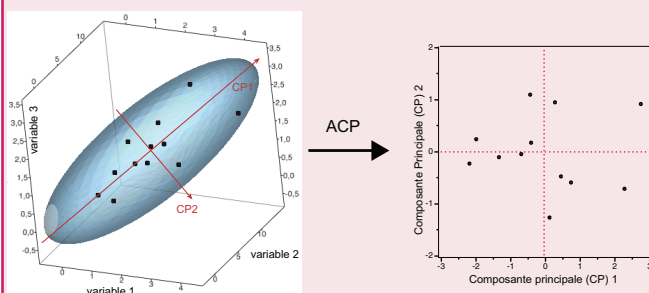


Figure A - Analyse en composantes principales, i.e. non supervisée (illustrée pour N = 3).

L'analyse en composantes principales (ACP) consiste à appliquer une rotation dans l'espace des variables à N dimensions, de façon à ce que le nouveau système d'axes, constitué par les composantes principales, maximise l'étalement des projections des individus (de coordonnées initiales définies par les colonnes de la matrice X). La première composante principale CP1 représente la direction de l'espace contenant la plus grande variance exprimée dans les données analysées. La deuxième composante principale CP2 représente la 2<sup>e</sup> direction de plus grande variance, dans le sous-espace orthogonal à CP1, etc. La procédure se poursuit jusqu'à ce que la totalité de la variance soit expliquée et permet ainsi de synthétiser l'information essentielle contenue au sein du jeu de données par un nombre limité de CP (généralement  $\ll N$ ). Chaque CP correspond à une combinaison linéaire des N variables d'origine, les poids représentant la

contribution des variables (métabolites) à ces composantes. La représentation de ces composantes permet d'en visualiser les signatures métaboliques associées (schéma des « loadings », voir figure 3 p. 62).

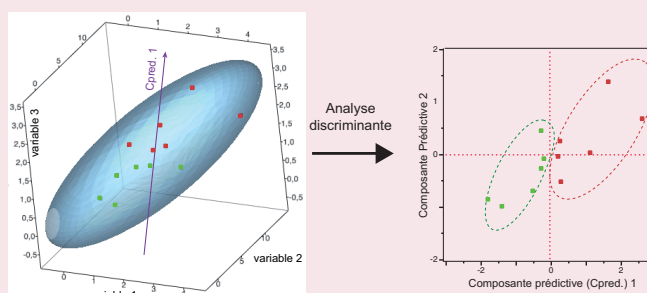


Figure B - Principe d'analyse supervisée discriminante (pour N = 3).

Les méthodes d'analyse supervisées (de type PLS) prennent en compte l'appartenance de chaque échantillon à un groupe donné. Le nouveau système d'axes (combinaisons linéaires des variables d'origine) maximise les distances inter-groupes. Les méthodes supervisées doivent être appliquées avec prudence, car elles peuvent conduire à des résultats surestimés. Chaque modèle statistique supervisé doit être rigoureusement évalué selon différentes méthodes, dont la plus connue est la validation croisée, pour en déterminer la *sensibilité* et la *prédictivité*. L'article de Broadhurst et Kell approfondit plus spécifiquement ces problématiques de validation [18].



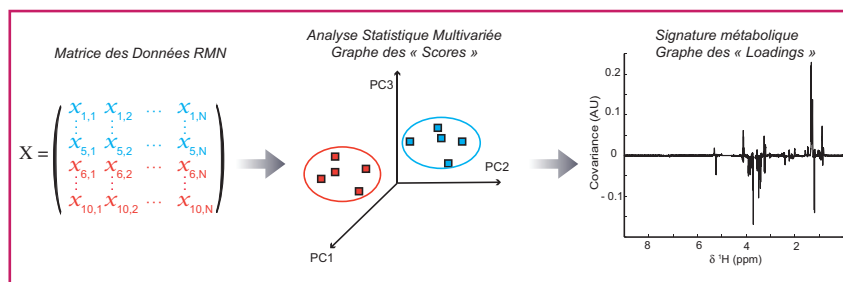


Figure 3 - Principe de modélisation statistique multivariée des données RMN.

particulièrement utile pour l'exploration des données et la détection d'échantillons aberrants. Elle permet de résumer la variance contenue dans les données dans un nombre réduit de composantes principales. Des algorithmes de classification hiérarchique ou « clustering » sont également largement utilisés pour une mise en évidence sans *a priori* de similarités en groupes d'échantillons. Cependant, il est possible que les sources principales de variance au sein d'une cohorte d'échantillons ne soient pas liées à l'effet étudié et par conséquent qu'une analyse de type ACP ne conduise pas à la séparation des échantillons souhaitée. Les méthodes supervisées telles que la régression des moindres carrés partiels (PLS, « partial least square »), l'analyse discriminante PLS ou l'analyse linéaire discriminante permettent alors d'extraire les composantes du profil métabolique spécifiquement corrélées aux paramètres d'intérêt de l'étude. Une extension de ces approches, la méthode OPLS (« orthogonalised projection on latent structures ») [2], développée plus récemment et qui sépare la variation corrélée au paramètre prédictif de classification de la variation orthogonale (non corrélée), est maintenant régulièrement employée pour l'analyse métabolomique.

## Vers la biologie des systèmes...

Pour une meilleure compréhension des mécanismes biologiques au sein d'approches de biologie intégrée des systèmes, la métabolomique peut être combinée à des données biochimiques ou à d'autres données « omiques », telles que la transcriptomique et la protéomique. Pour comprendre le métabolisme de la cellule et les réponses de ces cellules à des facteurs environnementaux et appréhender la complexité des systèmes biologiques, il est nécessaire d'étudier la cellule à différents niveaux, celui des gènes (génomique), de l'ARN messager (transcriptomique), des protéines (protéomique) et des métabolites (métabolomique), ainsi que leurs interactions (voir encadré 4). Dans la plupart des cas, ces différents types de données sont analysés de manière isolée, d'où une perte de l'information. Il existe différentes techniques statistiques permettant d'intégrer plusieurs jeux de données issues de techniques analytiques ou « omiques » différentes ; l'une des méthodes utilisées est l'O2PLS. Cette méthode supervisée est une généralisation de l'OPLS, décrite précédemment, qui permet la modélisation et la prédiction selon deux directions, à partir de matrices de données X et Y, représentant respectivement deux ensembles distincts de mesure, par exemple métabolomique par RMN et protéomique, collectés en parallèle sur des échantillons des mêmes individus. Ce type d'approche permet par exemple de déterminer les corrélations entre les métabolites discriminants et les protéines et vice versa [3].

## Reconstruction des réseaux métaboliques

La reconstruction des réseaux métaboliques à partir des données de métabolomique est une étape importante pour l'interprétation biologique des perturbations observées, car elle permet de visualiser les parties du réseau métabolique qui sont affectées par la perturbation biologique étudiée (pathologie, mutation, etc.). Les voies métaboliques perturbées impliquant les métabolites identifiés et

quantifiés par RMN sont repérées grâce à l'exploitation de bases de données comme KEGG (Kyoto Encyclopedia of Genes and Genomes), puis reconstruites et visualisées à l'aide d'un outil logiciel tel que cytoscape [4]. L'étude de Zhang *et coll.* illustre par exemple les perturbations métaboliques observées en fonction de l'inactivation du cycle de l'acide tricarboxylique (figure 4) [5].

### Encadré 4

#### D'autres grandes approches « omiques »...

**La génomique** est l'étude du génome, c'est-à-dire l'ensemble du matériel génétique d'un individu ou d'une espèce codée dans son ADN.

**La protéomique** désigne la science qui étudie des ensembles de protéines synthétisées par une cellule, leurs rôles, leurs structures, leurs localisations et leurs interactions. La démarche classique de l'analyse protéomique consiste, dans un premier temps, à séparer les protéines présentes dans une cellule ou un compartiment cellulaire. Chaque protéine d'intérêt est ensuite hydrolysée en fragments (peptides), qui sont analysés par spectrométrie de masse.

**La transcriptomique** est l'étude de l'ensemble des ARN messagers produits lors du processus de transcription d'un génome. Elle repose sur la quantification systématique de ces ARNm, ce qui permet d'avoir une indication relative du taux de transcription de différents gènes dans des conditions données. Plusieurs techniques, telles que les puces à ADN ou la PCR quantitative, donnent accès à ce type d'informations.

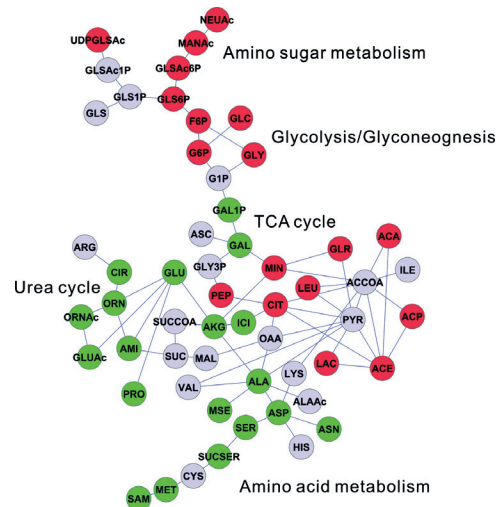


Figure 4 - Réseau métabolique représentant les variations des concentrations des métabolites causées par l'inactivation du cycle de l'acide tricarboxylique (TCA) : sont représentés en rouge ou vert les métabolites dont la concentration augmente ou respectivement diminue. Reproduit de [5] avec permission, © 2011 American Chemical Society.

La suite de cet article s'attache à présenter quelques grands domaines d'applications de l'analyse métabolomique par RMN concernant l'étude des plantes, la toxicologie ou la médecine, au travers de quelques exemples choisis.

## Applications à l'analyse par RMN métabolomique des plantes

À l'heure actuelle, environ 23 000 métabolites ont été identifiés dans les plantes (toutes variétés confondues), mais il est cependant estimé que ce royaume compterait entre 200 000 à 1 million de composés [6]. L'analyse métabolomique de tels systèmes complexes peut non seulement fournir des informations sur la nature fondamentale du phénotype des plantes en relation avec leur développement, leur physiologie et l'environnement, mais aussi aider à une meilleure compréhension de la nature interactive des réseaux métaboliques et de leurs réponses face aux changements génétiques. Bien que l'utilisation de la RMN pour l'analyse de plantes reste limitée, la plupart des métabolites ayant une concentration plus faible que la limite de détection de l'instrument, elle offre toutefois un certain nombre d'avantages par rapport aux autres techniques analytiques, tels qu'une préparation des échantillons simple et rapide, un temps d'acquisition spectrale très court (moins de 10 min en moyenne) et une excellente reproductibilité. De plus, la RMN permet de visualiser et de quantifier sur un même spectre les métabolites primaires (sucres, acides aminés...), majoritaires dans les plantes, et les métabolites secondaires (flavonoïdes, glucosinolates...) en fortes concentrations dans certains organismes.

Une particularité des plantes est leur très forte sensibilité aux variations environnementales (température, lumière, humidité), ce qui rend difficile la mise au point d'un protocole d'analyse reproductible. Il est souvent nécessaire de rassembler plusieurs plantes pour former un échantillon, pour des raisons de quantité de matériel, mais également pour moyenniser cette variabilité biologique. De nombreux efforts ont été faits pour mettre au point des procédés d'extraction fiables, capables de fournir une représentation fidèle du métabolisme [7]. Dans la suite, est présentée une sélection des applications de la RMN métabolomique pour l'analyse des plantes, dans deux des domaines les plus étudiés : les organismes génétiquement modifiés (OGM) et les interactions plantes/micro-organismes.

### Organismes génétiquement modifiés

La modification génétique des cultures agricoles (blé, riz, maïs...) a souvent pour but d'améliorer leur résistance aux agressions environnementales (climat, insectes) ou leur valeur nutritionnelle (augmentation de la concentration en antioxydants). Ces modifications génétiques consistent en la surexpression, l'addition, la suppression ou l'inactivation d'un ou de plusieurs gènes, spécifiquement sélectionnés pour leur influence sur le métabolisme de la plante. Cependant, les interactions entre métabolites sont difficiles à prévoir, et aux changements métaboliques attendus, peuvent s'associer des variations inattendues, pouvant rendre la plante impropre à la consommation. La recherche de méthodes capables d'évaluer ces risques de façon fiable est donc une problématique essentielle. La métabolomique par RMN, méthode quantitative et non ciblée, représente un outil de choix pour détecter ces modifications métaboliques inattendues. Elle permet de comparer l'ensemble du métabolisme des extraits de plantes non modifiées par rapport à leurs équivalents

transgéniques, dès que les molécules sont en concentration suffisante pour être détectées.

Considérons l'exemple du gène *Cry1Ab*, identifié comme étant responsable des qualités insecticides de la bactérie *Bacillus thuringiensis* qui est parfois utilisée comme insecticide en lieu et place d'un traitement chimique. L'expression de ce gène dans une plante, et notamment dans le maïs, permet la production par la plante elle-même d'une protéine insecticide dite protéine Crystal (la protéine *Cry1Ab*) qui tue en particulier les lépidoptères (papillons). La plupart des cultures agricoles génétiquement modifiées commercialisées sont porteuses de ce gène, et notamment les graines de maïs MON 810 commercialisées par Monsanto. Une équipe de recherche s'est intéressée aux variations métaboliques existant entre des graines de maïs exprimant ce gène par rapport aux mêmes graines non modifiées [8]. Dans leur étude, l'analyse statistique des spectres RMN du proton d'extraits hydroalcooliques des graines de maïs fournies par Monsanto a montré que la modification génétique entraînait des variations métaboliques importantes impliquant principalement l'asparagine, la glutamine, l'acide gamma-aminobutyrique, la choline et le succinate. Les auteurs ne concluent cependant pas sur les risques potentiels que ces modifications métaboliques peuvent engendrer en cas de consommation par l'homme ou l'animal. Ce sujet a cependant été abordé quelques années plus tard par une autre équipe de recherche, qui a étudié les effets allergènes de ces mêmes graines transgéniques (MON 810) sur des souris par RMN métabolomique [9]. Dans cette étude, les protéines extraites des graines de maïs transgéniques et de leur équivalent non-OGM ont été administrées à des souris par voie orale pendant 26 jours. L'urine de chaque souris a ensuite été prélevée avant consommation des protéines de maïs, et à 24 et 48 h après la dernière administration de celles-ci, puis analysée par RMN liquide. Une analyse discriminante linéaire a permis de mettre en évidence des différences métaboliques entre les souris traitées avec les protéines issues des graines OGM par rapport à celles traitées par les protéines issues des graines non-OGM (figure 5). Les auteurs de cette étude ont observé

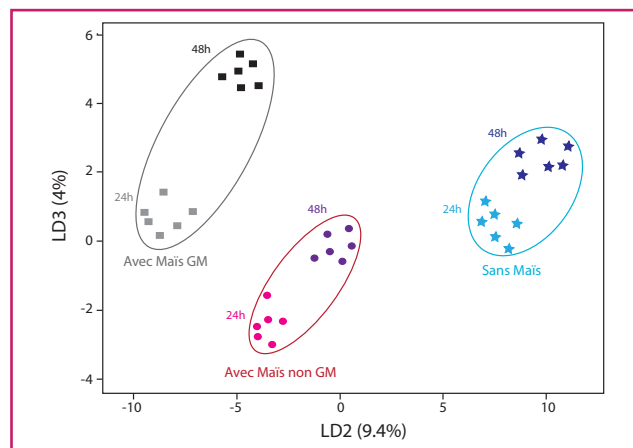


Figure 5 - Analyse linéaire discriminante des profils RMN  $^1\text{H}$  d'échantillons d'urine de souris prélevés pour des individus contrôles n'ayant pas consommé de protéines de maïs (étoiles), ayant consommé des protéines de maïs non transgéniques (ronds) ou transgéniques (carrés) pendant 26 jours.

Les trois différents groupes sont bien séparés suivant l'axe LD2, ce qui démontre l'effet du régime alimentaire administré sur le métabolisme des souris. Pour chaque groupe, les urines ont été prélevées 24 et 48 h après les 26 jours d'alimentation spécifique. Les variations métaboliques des urines au cours du temps sont bien mises en évidence le long de l'axe LD3.

une plus faible production de phénylalanine et d'hippurate et une plus forte production de tryptophane dans l'urine des souris traitées par les protéines du maïs MON 810. Cependant, ces variations ne représentent que 1 % de la variance métabolique totale et ne sont pas considérées comme des marqueurs d'inflammation ou de réactions immunologiques.

Ces deux exemples ne constituent qu'un petit échantillon des nombreuses études effectuées par RMN métabolomique sur les OGM. Des travaux plus fondamentaux sont également menés sur des plantes modèles (généralement sur la variété *Arabidopsis*) afin d'identifier la fonction métabolique de certains gènes. Ceci représente en principe la première étape dans la sélection des gènes qui pourraient être utilisés dans la modification génétique des cultures agricoles.

### Interactions entre les plantes et leur environnement

De la cohabitation entre les plantes et différents micro-organismes (pathogènes ou non pathogènes) résultent des variations importantes dans la réponse biochimique de la plante. On se trouve alors confronté à des réactions en cascades conduisant soit à une réponse mutuelle bénéfique, soit à un mécanisme de défense ciblé contre un pathogène. Par exemple, au moment de la pollinisation, les aracées vont sécréter des molécules volatiles et odorantes, imitant l'odeur du lieu de pondaison des mouches qui vivent dans leur environnement direct. Suivant la nature de ces mouches, ces plantes sécréteront donc de façon privilégiée une odeur de fruits pourris, d'excréments ou de charognes... Les mouches femelles ainsi leurrées viendront pondre dans les fleurs des aracées et repartiront avec le pollen de la plante qui pourra ainsi être dispersé. Ces mécanismes d'interaction, témoignant de la faculté d'adaptation des plantes à leur environnement, sont extrêmement intéressants, mais aussi excessivement complexes. De nombreux chercheurs se sont appliqués à comprendre ces mécanismes et certains ont focalisé leur intérêt sur l'identification des différentes classes chimiques impliquées dans les changements métaboliques liés à ces comportements, à l'aide de la RMN métabolomique.

Cette approche a été choisie par un groupe de chercheurs hollandais et français pour étudier les modifications du métabolisme des feuilles de tabac *Nicotiana tabacum* infectées par un virus (le virus mosaïque du tabac) [10]. Les auteurs de cette étude ont donc comparé les spectres RMN 2D J-résolue des feuilles infectées par rapport aux feuilles saines respectivement 1, 3, 7 et 10 jours après infection en utilisant une analyse en composantes principales. Ils ont ainsi montré que, dans une première étape, la plante développe un programme de mort cellulaire pour les cellules infectées par le virus. Métaboliquement, cette étape est associée à une altération des phénylpropanoïdes<sup>(1)</sup> impliqués dans le renforcement des parois cellulaires, ainsi qu'à une augmentation de la production en acide caffeoylquinique, connu pour contribuer à la résistance des plantes, et en acide alpha-linoléique, probablement impliqué dans la synthèse d'acide jasmonique qui contribue à l'activation des protéines de défense. Dans une seconde étape, ils assistent à une émission de signaux en cascades impliquant la production de composés tels que l'acide salicylique, qui permettent à la plante d'acquiescer une résistance systémique, et qui permettra à celle-ci de répondre de façon plus rapide et efficace en cas de futures expositions.

## Applications de la métabolomique en toxicologie

Les premières applications de la métabolomique dans le domaine de la toxicologie sont apparues il y a une quinzaine d'années. Les travaux fondateurs produits par l'équipe de J.K. Nicholson à Londres portant sur différentes études pharmacologiques ou toxicologiques ont permis de dépasser les limites rencontrées avec des techniques conventionnelles pour l'évaluation de la toxicité d'un certain nombre de substances modèles. L'objectif d'une étude métabolomique dans un contexte de toxicologie est de révéler un ensemble de biomarqueurs décrivant les perturbations métaboliques engendrées par l'exposition de l'organisme à des composés toxiques, permettant de produire des hypothèses quant au mécanisme de toxicité impliqué. De nombreuses études ont évalué les effets de substances toxiques par une approche de métabolomique par RMN, notamment pour l'évaluation toxicologique de médicaments, de préparations de médecine chinoise traditionnelle, de pesticides ou de contaminants environnementaux... Nous nous intéressons ci-après plus particulièrement à deux exemples dans le domaine de la pharmacologie et de la toxicologie de contaminants.

### Toxicologie des médicaments : pharmaco-métabolomique par RMN

L'évaluation de la toxicité des médicaments est un enjeu important pour l'industrie pharmaceutique. Les toxicités les plus problématiques sont celles qui apparaissent tard au cours d'un traitement, et qui ne sont pas détectables lors des études précliniques et cliniques de ces médicaments. Un défi majeur est donc l'identification précoce des biomarqueurs des réactions secondaires de ces médicaments, c'est-à-dire avant que le patient ne subisse de potentiels effets toxiques liés au traitement. L'analyse métabolomique par RMN de biofluides de rongeurs après traitement s'avère être une technique particulièrement pertinente pour l'évaluation de la toxicité de médicaments potentiels, démontrée par des études sur de nombreuses substances modèles. La prédiction de lésions au niveau du foie de rongeurs peut être également évaluée à partir d'études par RMN sur des biofluides prélevés en amont de tout traitement. Prenons l'exemple du paracétamol qui a été largement étudié par des approches de métabolomique par RMN.

Le paracétamol (connu également sous le nom d'acétaminophène) est un médicament couramment utilisé, analgésique et antipyrétique. Il est sans danger pour des doses thérapeutiques ; cependant, une overdose peut conduire à des lésions importantes du foie et du rein. La N-acétylcystéine (NAC) est l'un des antidotes les plus utilisés en cas d'intoxication par le paracétamol. Dans une première étude, Zira *et coll.* ont utilisé une approche métabolomique par RMN pour évaluer l'hépatotoxicité provoquée par une forte dose de paracétamol chez le lapin [11]. Les effets de l'antidote NAC ont également été étudiés. L'analyse en composantes principales (ACP) des données RMN a montré que l'intoxication par le paracétamol entraînait des variations métaboliques plasmatiques importantes (*figure 6*), impliquant une augmentation des lipides LDL et VLDL, des acides aminés, des lipides insaturés, ainsi qu'une diminution du glucose et du lactate. L'intoxication au paracétamol conduit donc à des variations du métabolisme des lipides et du

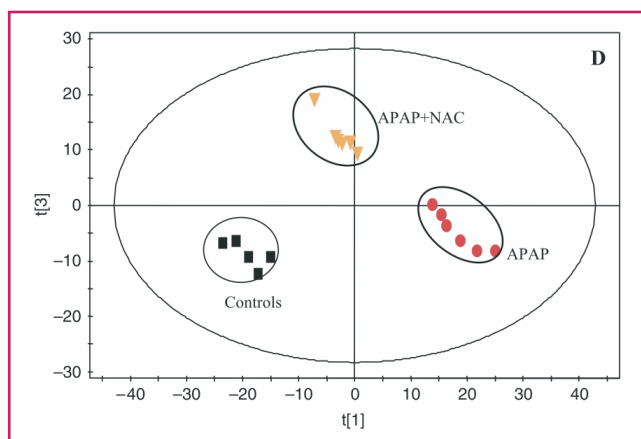


Figure 6 - Analyse en composantes principales de profils RMN  $^1\text{H}$  d'échantillons de sérum sanguin de lapin, prélevés pour des individus contrôles non traités (noir), traités par une forte dose de paracétamol, APAP 2 g/kg masse de l'animal (rouge), ou par une combinaison de paracétamol et d'antidote N-acétylcystéine, APAP 2 g/kg + NAC 150 mg/kg (jaune).

Cette analyse sépare distinctement les trois groupes d'échantillons, révélant des profils métaboliques caractéristiques de l'exposition au paracétamol ou à la combinaison paracétamol + antidote (reproduit de [11] avec permission).

métabolisme énergétique, ce qui a déjà été montré dans des études précédentes.

Le potentiel de la métabolomique pour prédire la toxicité d'un médicament au niveau du foie a été suggéré pour la première fois par Clayton *et coll.* [12]. Dans cette étude, des rats ont reçu une unique dose de paracétamol (600 mg/kg). Les urines avant le traitement (- 48 h, - 24 h) et après le traitement (24 h) ont été collectées et analysées par RMN du proton. Les lésions au niveau du foie ont été mesurées par histopathologie. Les rats traités ne présentent pas les mêmes lésions du foie. Les auteurs ont montré que les profils métaboliques enregistrés par RMN pour des urines collectées avant le traitement sont corrélés aux lésions du foie observées chez les rats après le traitement. Les principaux métabolites reliés au taux de nécrose dans le foie sont la taurine, le triméthylamine-N-oxyde (TMAO) et la bétaine. La toxicité du paracétamol est due à sa transformation en un métabolite réactif, la N-acétyl-p-benzoquinone (NAPQI) dans le foie, induisant les lésions hépatiques. La prédiction de l'hépatotoxicité due à l'intoxication au paracétamol par le métabolome avant le traitement peut s'expliquer par le fait que le métabolome avant traitement est corrélé à l'aptitude du foie à produire et accumuler ces réactifs. Cette technique est très prometteuse mais demande des développements supplémentaires pour pouvoir prédire de manière fiable la toxicité des médicaments chez l'homme.

### Toxicologie environnementale

Les perturbateurs endocriniens sont largement répandus dans l'environnement et dans la chaîne alimentaire, ce qui peut conduire à un risque pour la santé des consommateurs. Les polluants organiques permanents (POP) regroupent des substances telles que les polychlorobiphényles (PCB), les polychlorodibenzo-p-dioxines (PCDD), les polychlorobenzofuranes (PCDF) et des hydrocarbures aromatiques polycycliques (HAP). Ces composés sont très étudiés dans le domaine de la toxicologie environnementale car ce sont des substances toxiques, persistantes dans l'environnement et

qui s'accumulent dans les tissus vivants (bioaccumulation). La plupart des travaux de recherche sur la toxicité de ces composés et la santé humaine ont concerné des études toxicologiques conventionnelles de composés pris individuellement. Si l'on considère que les hommes sont exposés quotidiennement à des mélanges complexes de polluants, il apparaît important d'en étudier la toxicité, et la métabolomique semble être une méthode de choix pour cette évaluation.

Par exemple, Lu *et coll.* ont étudié récemment les effets des PCB et PCDD seuls ou en mélange chez des rats par une approche de métabolomique par RMN [13]. Les profils urinaires des rats ont été obtenus après administration orale de PCDD et PCB. En parallèle, l'histopathologie du foie et des analyses biochimiques du sérum ont montré une hépatotoxicité. L'analyse en composantes principales (ACP) des données RMN a montré que l'administration de ces contaminants entraînait des variations métaboliques importantes impliquant le lactate, le glucose, la taurine, la créatine, le 2-oxo-glutarate, le citrate, le succinate, l'hippurate et le triméthylamine-N-oxyde. Ces variations sont accentuées pour les rats traités avec le mélange des deux contaminants. Ces variations métaboliques peuvent être considérées comme des biomarqueurs potentiels de l'hépatotoxicité. Ces résultats ont montré que l'approche métabolomique par RMN est une technique efficace pour l'évaluation de la toxicité de mélanges de contaminants.

Une autre famille de contaminants alimentaires a été largement étudiée ces dernières années : les phtalates et le bisphénol A. Ces composés sont largement utilisés en plasturgie car ils permettent de renforcer la souplesse de ces matériaux. Les phtalates sont répertoriés comme de possibles anti-androgènes et possèderaient des propriétés œstrogéno-mimétiques. Dans une étude récente, les effets du diéthylhexylphtalate (DEHP) ont été étudiés par une approche métabolomique par RMN couplée à une approche transcriptomique [14]. Les souris ont été traitées tous les jours par gavage pendant 14 jours par différentes doses de DEHP (0, 30, 180 ou 1 100 mg/j/kg). Les urines, le plasma et les extraits aqueux de foie ont été analysés par RMN du proton. Les empreintes métaboliques ont montré un effet hypolipémiant du DEHP ainsi qu'une modification du métabolisme énergétique. En complément, une analyse du transcriptome a été réalisée sur les échantillons de foie des mêmes animaux. L'intégration des données de transcriptomique et métabolomique du foie à l'aide d'une analyse O2PLS a révélé une corrélation entre les effets du DEHP sur l'expression des gènes et sur les métabolites. L'analyse de type O2PLS met en évidence plusieurs covariations entre les transcrits et les métabolites affectés par le traitement au DEHP (figure 7). Par exemple, des variations de concentration des acides aminés (thréonine) sont reliées à des variations de l'expression de deux gènes (Vnn1 et Alas1). Ceci a conduit les auteurs à supposer que différentes régulations étaient impliquées dans l'augmentation de la synthèse de l'hème.

### Métabolomique par RMN et applications médicales

La RMN est recrutée depuis maintenant plusieurs années dans le cadre de projets de recherche translationnelle, vers des applications cliniques basées sur l'analyse métabolomique. Ces approches visent à l'utilisation de la spectrométrie de RMN comme outil à part entière de



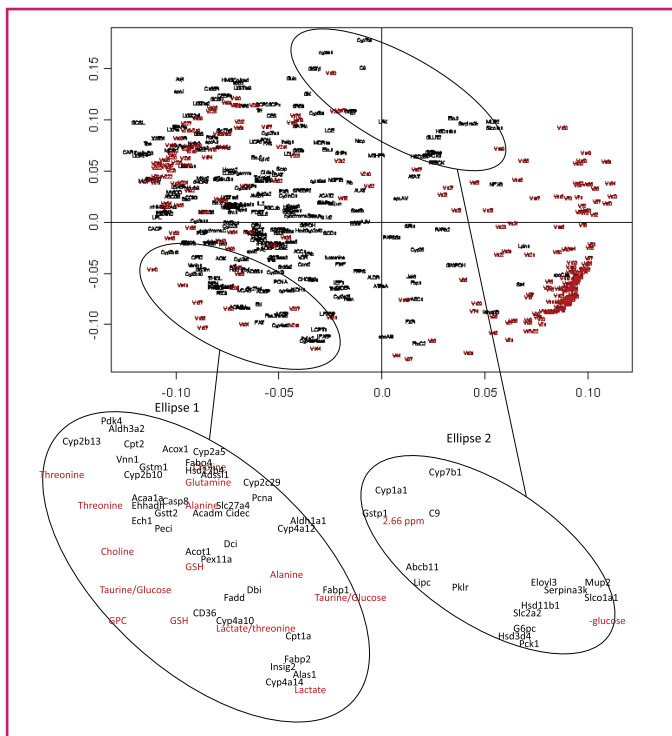


Figure 7 - Régression PLS des transcrits et des métabolites.

Les données de transcriptomique du foie et des métabolites des extraits aqueux du foie ont été traitées par une analyse O2PLS. Projection des métabolites (rouge) et des transcrits (noir). Les agrandissements (ellipses 1 et 2) montrent bien des corrélations entre les variations des métabolites et des transcrits. D'après [14] avec permission, © 2009 Elsevier.

diagnostic médical, permettant de détecter de façon non invasive, et potentiellement de manière précoce, des marqueurs métaboliques associés à diverses pathologies. De nombreuses études illustrent depuis une dizaine d'années les efforts et moyens mis en œuvre pour explorer le potentiel de cette approche pour l'étude de pathologies extrêmement variées – diabète, obésité, maladies cardio-vasculaires, cancer, maladies génétiques ou neurologiques, maladies infectieuses – au sein de projets cliniques localisés ou bien dans le cadre d'initiatives épidémiologiques plus larges. Au-delà de ces ambitions diagnostiques, la métabolomique se propose de sonder les effets de causalité et interactions entre nutrition, autres facteurs environnementaux (tabagisme, activité physique, etc.) et pathologies, et de mettre en évidence des populations ou individus à risque. Enfin, les approches de métabolomique représentent un potentiel intéressant pour la médecine personnalisée – par exemple pour l'adaptation du traitement à un patient selon des critères révélés par son profil métabolique urinaire ou sanguin – et prédictive, par la détermination précoce de l'évolution d'une pathologie (risques de complications, rechutes, effets toxiques d'un traitement...). Nous présentons plus spécifiquement ci-après quelques exemples d'applications de la métabolomique dans le domaine du cancer et de l'épidémiologie moléculaire.

### Métabolomique en cancérologie

De nombreux exemples dans la littérature traitent de la caractérisation par RMN des tissus tumoraux par la technique HR-MAS. Le potentiel de l'approche métabolomique a été en effet illustré en médecine en premier lieu pour l'analyse de biopsies cancéreuses, permettant de

démontrer d'une part la fonctionnalité de la technique comme outil de classification appliqué à la gradation des tumeurs (classification déjà maîtrisée, bien que de façon moins objective, par les techniques classiques d'histopathologie), et d'autre part de proposer une étude plus fine du métabolisme tumoral, se focalisant sur la compréhension des mécanismes et voies métaboliques impliqués dans la tumorigénèse ou la prolifération tumorale.

Plus récemment, la contribution de la métabolomique vers des applications cliniques en cancérologie s'est tournée vers l'étude non invasive des fluides excrétés ou circulants (urines et sérum ou plasma sanguins principalement) pour la détection de marqueurs métaboliques, ou *biomarqueurs*, associés aux pathologies cancéreuses. L'exemple suivant illustre plus spécifiquement l'une de ces initiatives prometteuses.

Une étude récente montre l'utilisation d'une analyse de métabolomique par RMN pour la détection des stades précoces I/II du carcinome épithélial de l'ovaire [15]. Cette étude exploratoire a été menée selon un design expérimental original utilisant de très faibles quantités d'échantillons (20 µl), à l'aide d'une sonde RMN à micro-flux continu (figure 8). L'utilisation de méthodes de régression logistique et d'analyse en composantes principales des spectres RMN  $^1\text{H}$  a permis de dévoiler une signature sérique spécifique associée aux patientes atteintes de cancer de l'ovaire, comparées à des individus sains. Elle a également montré la spécificité de cette signature par rapport à d'autres types de cancers, les patientes atteintes de cancer du rein avancé n'étant pas selon ce critère diagnostiquées de façon erronée comme atteintes d'un cancer de l'ovaire. Cette étude souligne donc le potentiel remarquable de la RMN comme outil de diagnostic non ou peu invasif en cancérologie, pouvant être en théorie appliqué à grande échelle pour la détection précoce, en particulier pour des populations à fort risque de cancer de l'ovaire. D'autres travaux de même

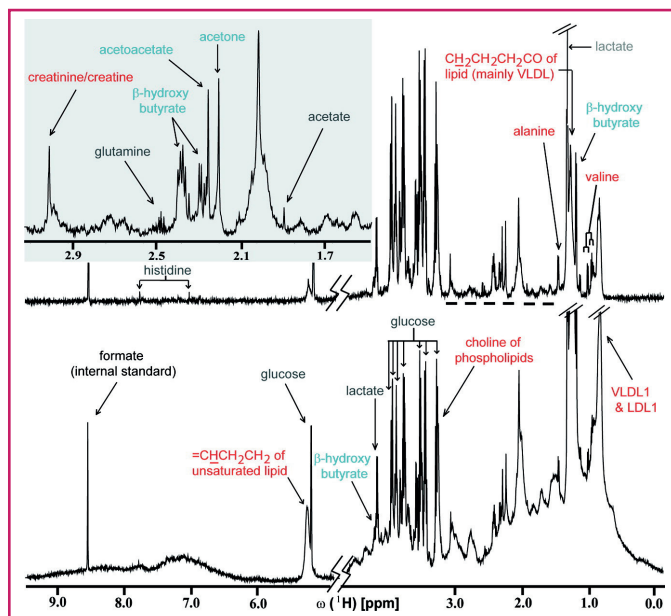


Figure 8 - Spectres  $^1\text{H}$  représentatifs de sérum sanguin de type CPMG (haut) ou d'acquisition classique (bas), pour une patiente atteinte de cancer épithélial de l'ovaire.

Les métabolites identifiés en rouge correspondent à des concentrations détectées significativement plus faibles pour des patientes cancéreuses par rapport à des sujets sains, et en bleu ceux qui sont présents en quantité plus élevée dans les cas de cancer. Reproduit de [15] avec permission, © 2011 American Chemical Society.

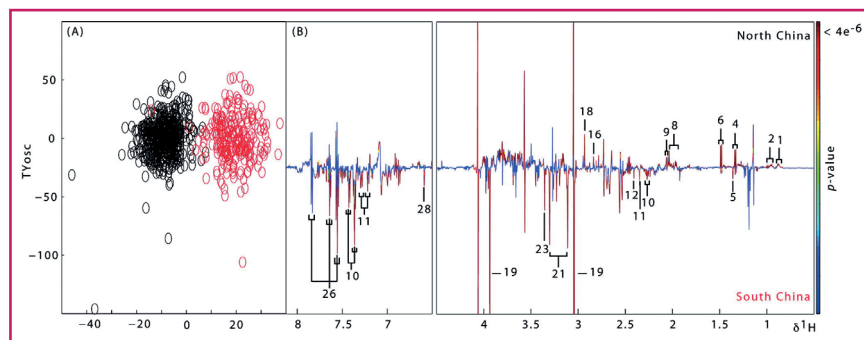


Figure 9 - (A) Diagramme des « scores » pour une analyse OPLS discriminante appliquée aux spectres RMN d'urine de populations du nord (noir) ou du sud (rouge) de la Chine, mettant en évidence une distinction nette entre les profils métaboliques observés pour ces deux populations. (B) Diagramme de covariance ou « loadings » identifiant les métabolites différenciant les populations nord-sud de la Chine.

Les signaux positifs représentent les métabolites observés en plus forte concentration pour la population du nord ; les massifs négatifs correspondent aux métabolites en concentration plus élevée pour la population du sud. 1 : acide pentanoïque/heptanoïque ; 2 : acides aminés branchés ; 4 : lactate ; 5 : 2-hydroxyisobutyrate ; 6 : alanine ; 8 : N-acétyls de fragments de glycoprotéines (incluant l'uromoduline) ; 9 : acide N-acétyl neuraminique ; 11 : sulfate de 4-crésyl ; 12 : succinate ; 16 : méthylguanidine ; 18 : diméthylglycine ; 19 : créatine ; 21 : prolinebétaine ; 23 : scyllo-inositol ; 10 : phénylacétylglutamine ; 26 : hippurate ; 28 : trans-aconitate (d'après [16] avec permission, © 2010 American Chemical Society).

nature démontrent des qualités similaires de l'analyse métabolomique par RMN pour la détection précoce ou de stades plus avancés dans différents types de cancers.

## Épidémiologie moléculaire

De nouvelles initiatives et un grand intérêt sont apparus depuis quelques années pour l'utilisation de techniques de profilage métabolique dans le cadre d'études épidémiologiques à grande échelle. La robustesse de la technique de RMN rend son utilisation dans ce cadre particulièrement pertinente, et la disponibilité ou la constitution de cohortes d'échantillons adaptées, dont les conditions de prélèvement et de conservation selon des protocoles rigoureux et standardisés constituent un élément clé pour la fiabilité des études de métabolomique, ont permis le développement de projets exploratoires d'épidémiologie moléculaire par RMN métabolomique.

Dans le cadre d'un projet à grande échelle baptisé INTERMAP (« International Study of Macro/Micronutrients and Blood Pressure »), Yap *et coll.* se sont intéressés aux différences pouvant être détectées dans des profils urinaires pour des populations des régions du nord ou du sud de la Chine, pour lesquelles les taux de pathologies cardiaques ou d'accidents vasculaires cérébraux constatés sont sensiblement différents [16]. L'étude par RMN d'échantillons d'urines pour plus de 750 volontaires provenant de ces deux régions a permis d'identifier par analyse discriminante de type OPLS une signature métabolique caractéristique distinguant ces deux populations (figure 9). Un ensemble de métabolites associés à ces différences a pu être identifié, comprenant notamment pour les populations du sud des co-métabolites de la flore microbienne (hippurate, 4-crésyl sulfate, phénylacétylglutamine, 2-hydroxyisobutyrate). Ces observations soulignent l'importance des facteurs environnementaux (régime alimentaire), du métabolisme endogène, ainsi que du co-métabolisme mammifère/flore intestinale, qui peuvent expliquer les différences nord-sud observées en Chine concernant les risques de survenue de maladies cardiovasculaires.

Dans une autre étude basée sur cette même initiative INTERMAP, Holmes *et coll.* ont étudié plus largement la

diversité du phénotype métabolique humain et son association avec le régime alimentaire et la tension artérielle [17]. Des échantillons d'urine ont été analysés pour plus de 4 500 individus provenant de différentes populations de Chine, du Japon, des États-Unis et du Royaume-Uni. Ils ont montré que les profils d'excrétion urinaire permettaient d'établir des phénotypes métaboliques différenciés pour des populations d'Asie orientale ou provenant de régions occidentales caractérisées par des habitudes alimentaires et des taux de maladies cardio-vasculaires très différents, de même que pouvaient être identifiés des sous-groupes correspondant à une alimentation végétarienne, riche en protéines animales, ou caractérisant des profils distincts de pression artérielle (figure 10). Ces profils métaboliques font apparaître différents métabolites discriminants, tels que le formate, l'alanine ou l'hippurate, reflétant le régime alimentaire

et l'activité de la flore microbienne (ou *microbiote*), et associés à l'hypertension artérielle.

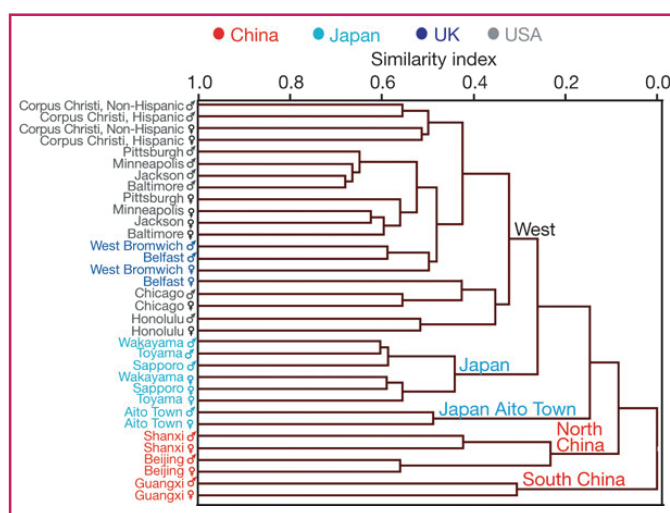


Figure 10 - Analyse par clustering hiérarchique pour des échantillons humains d'urine de populations variées analysés par RMN 1D  $^1\text{H}$ .

Le dendrogramme souligne les similarités/différences entre populations, la proximité des différentes classes dans ce diagramme étant directement reliée à la similitude des profils métaboliques moyens pour ces populations. L'indice de similarité en abscisse mesure la distance multivariée entre clusters d'échantillons (une valeur nulle indique une séparation maximale, une valeur unité correspond à une identité des profils métaboliques caractéristiques des sous-populations concernées). Certaines populations locales présentent des empreintes métaboliques remarquablement singulières, comme la population de la province de Guangxi en Chine. Les profils métaboliques des habitants de la ville rurale d'Aito au Japon se démarquent également de ceux de populations plus urbaines du Japon (Wakayama, Toyama, Sapporo), qui présentent des profils nettement plus similaires à ceux observés pour des populations occidentales. Reproduit de [17] avec la permission de Macmillan Publishers Ltd, © 2008.

Les quelques exemples développés ci-dessus donnent un aperçu du potentiel analytique et de la diversité des applications offertes par les approches de métabolomique par RMN. Il est à noter que ces applications ne se limitent pas aux domaines illustrés dans cette revue, mais couvrent un champ beaucoup plus large, sur le plan fondamental par exemple au sein d'études de biologie cellulaire ou de génomique fonctionnelle, ou bien vers des applications

industrielles de contrôle-qualité alimentaire, de lutte contre la fraude concernant les produits d'origine contrôlée, ou encore de lutte contre le dopage.

## Note et références

- (1) *Phénylpropanoïdes* : classe de composés organiques, dérivés de plantes, biosynthétisés à partir d'un acide aminé, la phénylalanine. Ils ont une grande variété de fonctions, comme la défense contre les herbivores, les attaques microbiennes et les autres agresseurs potentiels.
- [1] Lindon J.C., Nicholson J.K., Recent advances in high-resolution NMR spectroscopic methods in bioanalytical chemistry, *TrAC-Trends Anal. Chem.*, **1997**, 16(4), p. 190.
- [2] Trygg J., Wold S., Orthogonal projections to latent structures (O-PLS), *J. Chemom.*, **2002**, 16(3), p. 119.
- [3] Rantalainen M., Cloarec O., Beckonert O., Wilson I.D., Jackson D., Tonge R., Rowlinson R., Rayner S., Nickson J., Wilkinson R.W., Mills J.D., Trygg J., Nicholson J.K., Holmes E., Statistically integrated metabolomic-proteomic studies on a human prostate cancer xenograft model in mice, *J. Proteome Res.*, **2006**, 5(10), p. 2642.
- [4] www.cytoscape.org.
- [5] Zhang B., Halouska S., Schiaffo C.E., Sadykov M.R., Somerville G.A., Powers R., NMR analysis of a stress response metabolic signaling network, *J. Proteome Res.*, **2011**, 10(8), p. 3743.
- [6] Saito K., Matsuda F., Metabolomics for functional genomics, systems biology, and biotechnology, *Annual Review of Plant Biology*, **2010**, 61, p. 463.
- [7] Kim H.K., Choi Y.H., Verpoorte R., NMR-based metabolomic analysis of plants, *Nat. Protocols*, **2010**, 5(3), p. 536.
- [8] Manetti C., Bianchetti C., Casciani L., Castro C., Di Cocco M.E., Miccheli A., Motto M., Conti F., A metabolomic study of transgenic maize (*Zea mays*) seeds revealed variations in osmolytes and branched amino acids, *J. Exp. Bot.*, **2006**, 57(11), p. 2613.
- [9] Adel-Patient K., Guimaraes V.D., Paris A., Drumare M.-F., Ah-Leung S., Lamourette P., Nevers M.-C., Canlet C., Molina J., Bernard H., Creminon C., Wal J.-M., Immunological and metabolomic impacts of administration of Cry1Ab protein and MON 810 maize in mouse, *PLoS One*, **2011**, 6(1), p. e16346.
- [10] Choi Y.H., Kim H.K., Linthorst H.J.M., Hollander J.G., Lefeber A.W.M., Erkelens C., Nuzillard J.M., Verpoorte R., NMR metabolomics to revisit the tobacco mosaic virus infection in *Nicotiana tabacum* leaves, *J. Nat. Prod.*, **2006**, 69(5), p. 742.
- [11] Zira A., Mikros E., Giannioti K., Galanopoulou P., Papalois A., Liapi C., Theocharis S., Acute liver acetaminophen toxicity in rabbits and the use of antidotes: a metabolomic approach in serum, *J. Appl. Toxicol.*, **2009**, 29(5), p. 395.
- [12] Clayton T.A., Lindon J.C., Cloarec O., Antti H., Charuel C., Hanton G., Provost J.P., Le Net J.L., Baker D., Walley R.J., Everett J.R., Nicholson J.K., Pharmaco-metabolomic phenotyping and personalized drug treatment, *Nature*, **2006**, 440(7087), p. 1073.
- [13] Lu C., Wang Y., Sheng Z., Liu G., Fu Z., Zhao J., Zhao J., Yan X., Zhu B., Peng S., NMR-based metabolomic analysis of the hepatotoxicity induced by combined exposure to PCBs and TCDD in rats, *Toxicol. Appl. Pharmacol.*, **2010**, 248(3), p. 178.
- [14] Eveillard A., Lasserre F., de Tayrac M., Polizzi A., Claus S., Canlet C., Mselli-Lakhal L., Gotardi G., Paris A., Guillou H., Martin P.G.P., Pineau T., Identification of potential mechanisms of toxicity after di-(2-ethylhexyl)-phthalate (DEHP) adult exposure in the liver using a systems biology approach, *Toxicol. Appl. Pharmacol.*, **2009**, 236(3), p. 282.
- [15] Garcia E., Andrews C., Hua J., Kim H.L., Sukumaran D.K., Szyperski T., Odunsi K., Diagnosis of early stage ovarian cancer by H-1 NMR metabolomics of serum explored by use of a microflow NMR probe, *J. Proteome Res.*, **2011**, 10(4), p. 1765.
- [16] Yap I.K.S., Brown I.J., Chan Q., Wijeyasekera A., Garcia-Perez I., Bictash M., Loo R.L., Chadeau-Hyam M., Ebbels T., Iorio M.D., Maibaum E., Zhao L., Kesteloot H., Daviglus M.L., Stamler J., Nicholson J.K., Elliott P., Holmes E., Metabolome-wide association study identifies multiple biomarkers that discriminate North and South Chinese populations at differing risks of cardiovascular disease: INTERMAP study, *J. Proteome Res.*, **2010**, 9(12), p. 6647.
- [17] Holmes E., Loo R.L., Stamler J., Bictash M., Yap I.K.S., Chan Q., Ebbels T., De Iorio M., Brown I.J., Veselkov K.A., Daviglus M.L., Kesteloot H., Ueshima H., Zhao L., Nicholson J.K., Elliott P., Human metabolic phenotype diversity and its association with diet and blood pressure, *Nature*, **2008**, 453(7193), p. 396.
- [18] Broadhurst D.I., Kell D.B., Statistical strategies for avoiding false discoveries in metabolomics and related experiments, *Metabolomics*, **2006**, 2(4), p. 171.



C. Canlet



L. Shintu



B. Elena-Herrmann

### Cécile Canlet

est ingénieur de recherche en RMN de la plate-forme AXIOM-MetaToul, INRA UMR 1331, TOXALIM, Université de Toulouse<sup>1</sup>.

### Laetitia Shintu

est maître de conférences à l'Institut des Sciences Moléculaires de Marseille<sup>2</sup>.

### Bénédicte Elena-Herrmann (auteur correspondant)

est ingénieur de recherche au CNRS à l'Institut des Sciences Analytiques, Centre de RMN à Très Hauts Champs, Université de Lyon<sup>3</sup>.

<sup>1</sup> INRA UMR 1331, TOXALIM, Université de Toulouse, ENVT, INP, UPS, 180 chemin de Tournefeuille, F-31027 Toulouse.

Courriel : cecile.canlet@toulouse.inra.fr

<sup>2</sup> Institut des Sciences Moléculaires de Marseille, Aix-Marseille Université, iSm2-UMR CNRS 7313, Campus scientifique de Saint Jérôme, case 512, F-13397 Marseille Cedex 20.

Courriel : laetitia.shintu@univ-amu.fr

<sup>3</sup> Université de Lyon, CNRS/École Normale Supérieure de Lyon/UCBL, Institut des Sciences Analytiques, Centre de RMN à Très Hauts Champs, 5 rue de la Doua, F-69100 Villeurbanne.

Courriel : benedict.elena@ens-lyon.fr



Société Chimique de France

facebook

**La SCF sur Facebook, vous aimez ?**  
Parlez-en autour de vous,  
et invitez vos amis et collègues à nous rejoindre !

• [www.facebook.com/pages/Soci%C3%A9t%C3%A9-Chimique-de-France/114534205270205](http://www.facebook.com/pages/Soci%C3%A9t%C3%A9-Chimique-de-France/114534205270205)